

SwissBib: ein Metakatalog NextGeneration oder 2.0

Von 2008 bis 2011 baut das Projekt SwissBib (Der Fortschritt des Projekts wird in einem Projektwiki zusammen mit der Planung für die zweite Projektphase dokumentiert: <http://www.swissbib.org>. Der Zugang zur Suchmaschine erfolgt über <http://www.swissbib.ch>. Dieser Zugang ist bis zum Start der öffentlichen Betaphase passwortgeschützt. [Stand: 30.6.2009]) im Rahmen des Schweizer Förderprogramms E-Lib.ch einen gesamtschweizerischen Metakatalog für die Verbände der Schweizer Universitätsbibliotheken und der Schweizerischen Nationalbibliothek auf. Weiterhin werden Metadaten von Institutional Repositories und von E-Archiving-Projekten mit einbezogen sowie Daten aus den Kataloganreicherungsprojekten der Universitätsbibliotheken.

Ziel des Projekts unter der Führung der Universitätsbibliothek Basel ist es, den Metakatalog basierend auf aktuellen Technologien zu realisieren, das heisst mit OPAC-2.0-Technologien. Im Rahmen des Projekts werden acht Server von 13 Teilverbänden und der Nationalbibliothek einbezogen – insgesamt 17 Mio. bibliographische Datensätze, acht Normdatenbanken für Personennamen und Körperschaften, eine Vielzahl von Schlagwortsystemen und Klassifikationen. Aber auch unterschiedliche Ausleihsysteme, Benutzerverwaltungen, Betriebskulturen und unterschiedliche Grade an gegenseitiger Integration und Zusammenarbeit müssen einbezogen werden.

Der Umfang des Projekts macht die Aufteilung in mehrere Arbeitspakete unumgänglich. So wurde entschieden, es in zwei Hauptphasen zu gliedern. Die erste Phase bis Ende 2009 widmet sich der Evaluation möglicher Systeme, dem Kauf der bestgeeigneten Software und dem Aufbau einer öffentlich benutzbaren Version. Die zweite Phase von 2010 bis 2011 widmet sich der Verbesserung und dem Ausbau des Service und greift für die Recherchequalität zentrale Themen wie Mehrsprachigkeit und die Einbindung externer Normdaten via VIAF und MACS auf. Das Projekt VIAF – virtual international authority file beschäftigt sich mit der Zusammenführung verschiedener Autoritätsdatenbanken zu einer Konkordanz: <http://www.viaf.org> [Stand: 30.6.2009]. Das Projekt MACS – multilingual access to subjects beschäftigt sich mit der Zusammenführung verschiedener Sacherschliessungssprachen zu einer Konkordanz: <https://macs.hoppie.nl/pub> [Stand: 30.6.2009].

Ein erster funktionaler Prototyp ging Ende April 2009 online und wird durch die SwissBib-Partner getestet.

Auf Ende September 2009 sollen die angeregten Verbesserungen eingepflegt und in einer internen Betaversion für interessierte Partnerinstitutionen geöffnet werden. Auf Ende 2009 wird SwissBib für die breite Öffentlichkeit sichtbar werden.

Modelle für SwissBib: OPAC 2.0 und «federated search»

Spätestens mit dem Artikel von Karen G. Schneider^{Es handelt sich um einen dreiteiligen Blogbeitrag, stellvertretend sei der erste Teil zitiert: Karen G. Schneider, How OPACs Suck, Part 1: Relevance Rank (Or the Lack of It), in: Ala Tech Source 13.3.2006.}

<http://www.techsource.ala.org/blog/2006/03/how-opacs-suck-part-1-relevance-rank-or-the-lack-of-it.html>

[Stand: 30.6.2009] ist ein zentraler Faktor für die Entwicklung benannt, die die Bibliothekswelt seit mehreren Jahren beschäftigt: Schwindende Nutzerzahlen bei den Onlinekatalogen und auf Bibliotheksseiten generell. Die Suchapplikationen der Bibliothekssysteme gehen nicht mehr auf die Bedürfnisse der meisten Bibliotheksbenutzer ein, die sich anderen Quellen zugewandt haben und von diesen in ihren Bedürfnissen massgebend geprägt werden. Eindrücklich dokumentiert dies die OCLC-Studie zur Wahrnehmung von Bibliotheken und weiteren Informationsmitteln durch Studierende: OCLC, College Students' Perceptions of Libraries and Information Resources. A Report to the OCLC Membership, Dublin Ohio 2006. <http://www.oclc.org/reports/pd...> [Stand: 30.6.2009]. Stellvertretend für eine Vielzahl von Untersuchungen zum veränderten Nutzerverhalten sei hier eine Studie des britischen CIBER erwähnt, die auf gross angelegter Log-Datenanalyse beruht: Ian Rowlands et al., The Google generation; the information behaviour of the researcher of the future, in: Aslib Proceedings Vol. 60, No. 4 (2008), S. 290–310.

Die Reaktion auf die vehement vorgetragene Kritik an den bestehenden Systemen aus der bibliothekarischen Zunft ist nicht ausgeblieben. Neben einer Vielzahl von Untersuchungen über die Bedürfnisse von Nutzern hat sich vor allem im Bereich der Suchapplikationen einiges getan. Die kommerziellen Anbieter von Bibliothekssystemen widmen sich wieder vermehrt der Suche, und eine lebendige OpenSource-Gemeinde entwickelt neue OPAC – eben NextGeneration oder auch 2.0. Eine Auflistung aktueller Suchapplikationen ist zu finden unter: http://www.swissbib.org/wiki/index.php?title=Software_solutions [Stand: 30.6.2009]. Anne Christensen, Projektverantwortliche des Hamburger Katalog-2.0-Projekts beluga, hat eine Karte mit europäischen OPAC-2.0-Projekten online gestellt: <http://www.communitywalk.com/m...> [Stand: 30.6.2009].

Grundsätzlich ist diesen Systemen gemein, dass sie auf moderner Suchmaschinentechnologie aufbauen und mit dieser versuchen, möglichst alle Informationen einer Bibliothek über eine Suchoberfläche verfügbar zu machen. Darunter ist technologisch zu verstehen, dass die bisherigen direkt auf den Suchmöglichkeiten der Datenbank des Bibliothekssystems realisierten OPAC durch von diesen unabhängige, auf die Suche in grossen und heterogenen Datensätzen spezialisierte Suchmaschinen ersetzt werden. Dabei orientiert man sich funktional an den Möglichkeiten, wie man sie von Suchmaschinenanbietern wie Google, Yahoo und fachlich spezialisiert von WorldCat sowie Internetversandhäusern wie Amazon kennt. Anne Christensen bietet eine Auflistung der grundlegenden Funktionen eines OPAC 2.0 in: Anne Christensen, Partizipative Entwicklung von Diensten in der Bibliothek 2.0: Methoden und Ergebnisse aus Katalog-2.0-Projekten, in: Bibliotheksdienst 43. Jg. (2009), S. 527–537, S. 529.

Während sich die technische Umsetzung wie auch die Vorstellungen hinsichtlich der Benutzerführung stark verändert haben, ist im Bereich der Daten die grundlegende Ausrichtung dieselbe geblieben. Weiterhin wird davon ausgegangen, dass der Hauptteil der zu durchsuchenden und zu präsentierenden Daten aus einer Quelle stammt – nämlich der bibliografischen Datenbank der Bibliothekssysteme. Dies hat Auswirkungen auf die Art und Weise, wie Daten aufbereitet werden, und reicht von einer mehr oder weniger ausgefeilten Technik mit Mechanismen zur Datenedublierung und -anreicherung bis zum kompletten Weglassen solcher Funktionen.

Die zweite, ältere und insgesamt pragmatischere Lösung ist in der klassischen «federated search» zu sehen, die sich insgesamt hoher Verbreitung erfreut und in der Schweiz mit der IDS-Gesamtabfrage, RERO recherche fédérée und CHVK gleich mehrfach mit ähnlicher Zielsetzung eingesetzt wird. Diese Systeme basieren darauf, dass Zielsysteme in dem Moment abgefragt werden, in dem ein Benutzer seine Suchanfrage absetzt. Die Resultate werden aufbereitet, sobald sie geliefert werden. Damit solche Systeme gut funktionieren, müssen alle abgefragten Zielsysteme schnell antworten, und die Treffermenge sollte überschaubar sein. Ist dies nicht der Fall, reagiert das System für den Benutzer langsam und liefert aus Performancegründen nur Teiltreffermengen. Das heisst, jede nachgelagerte Datenaufbereitung, wie beispielsweise die Dedublierung und Facettierung durch Metalib ^{Metalib} ist das federated-search-Suchportal von Ex Libris, das in der Schweiz vom IDS und von RERO eingesetzt wird., ist dann unvollständig und tendenziell für die Benutzer verwirrend. Ihr Vorteil ist aber, dass vor allem in heterogenen Umgebungen die Suchanfragen von den optimal auf die eigenen Daten abgestimmten Lokalsystemen verarbeitet werden und zentral wenig Aufwand getrieben werden muss.

Der Metakatalog SwissBib im Spannungsfeld zwischen OPAC 2.0 und «federated search»

Betrachtet man nun die Probleme des herkömmlichen Ansatzes und kontrastiert sie mit den Systemen der neuen Generation, bekommt man eine gute Grundlage für den Aufbau eines Metakataloges wie SwissBib. Da mit Blick auf moderne Suchmaschinen die Geschwindigkeit des Service zentral ist, fällt die Option «federated search» weg, wie sie mit Blick auf die Heterogenität der Daten und die Diversität der Schweizer Bibliothekslandschaft einleuchtend wäre. Vor diesem Hintergrund kommt der Frage der Datenhaltung eine zentrale Rolle zu. Besieht man sich die Daten und die Organisationsstruktur in der Schweiz, kann eine Homogenisierung der Daten nicht kollaborativ auf lokaler Ebene in nützlicher Frist erfolgen. Es bleibt als sinnvollste Lösung die Sammlung und zentrale Aufbereitung der Daten zum Zweck der Suche unabhängig von den Lokalsystemen der einzelnen Verbände und Bibliotheken. Das heisst, die bibliografischen Daten der einzelnen Verbände werden zentral gesammelt, und der so entstehende Datenpool regelmässig aktualisiert. Genau bei der Aufbereitung heterogener Daten haben die meisten neuen Systeme Schwächen, während SwissBib in diesem Bereich einen Schwerpunkt setzt. Neben dem Gewinn an Geschwindigkeit bei der Suche über einen so vereinigten Datenpool bietet dieses Vorgehen die grösstmögliche Flexibilität im Umgang mit den Daten. Dieser Datenpool bildet dann die Grundlage für die einheitliche Aufbereitung durch eine Suchmaschine für die Suche, wobei die zentrale Datenhaltung auf demselben System wie die Suchmaschine gegenüber der «federated search» einen massiven Geschwindigkeitsvorteil bietet.

Es muss berücksichtigt werden, dass eine qualitativ ansprechende Suche in höchstem Masse von der Homogenität des zugrunde liegenden Datenmaterials hinsichtlich Erfassungsstandards und Format und einer grösstmöglichen Informationsdichte abhängt. Dies hat direkten Einfluss auf die von der Suchmaschine bereitgestellte Funktionalität, vor allem bei Filtern und der Trefferaufbereitung. So werden mit dem Einsatz von Facetten auch Inkonsistenzen in den Daten auf die gleiche übersichtliche Weise herausgehoben wie die eigentlich gewünschten Suchbegriffe. Auch die Qualität des Rankings steigt, wenn Informationen nach einheitlichen Kriterien ausgewertet werden können. Fehler in diesem Bereich sind – wie die Studie von OCLC zu den Nutzerbedürfnissen gegenüber Suchmaschinen zeigt – für Benutzer besonders störend, da man dieser Funktion hohen Stellenwert einräumt und sich auf sie verlassen möchte [1](#).

Die SwissBib-Oberfläche in ihrer Funktionalität und Gestaltung orientiert sich an modernen Systemen¹ OCLC, Online Catalogs: What Users and Librarians Want. An OCLC Report, Dublin Ohio 2009, v.a. S. 11–21 und S. 51–52 (<http://www.oclc.org/reports/onlinecatalogs/fullreport.pdf> [Stand: 30.6.2009]) Auf die Diskussion zur Repräsentativität der Studie sei an dieser Stelle hingewiesen. In jedem Fall gibt die Studie durch ihre Konzentration auf WorldCat klare Hinweise zu den grundlegenden Fragen für die Ausrichtung eines Metakataloges. Hierbei ist es zentral, dass die vorhandenen Funktionen und Informationen in einer möglichst intuitiven Weise gruppiert und gestaltet sind. Der Ersteinstieg wird einfach gehalten sein, ohne aber die Möglichkeiten für fortgeschrittene Benutzer zu beschränken. Die Präsentation der Resultate soll so aussagekräftig als möglich sein, ohne allerdings überladen zu wirken. Wichtig ist in diesem Zusammenhang, dass dem Benutzer möglichst klar wird, ob es sich beim jeweiligen Treffer um eine Onlineressource, eine Monografie oder beispielsweise eine DVD handelt. Von enormer Wichtigkeit sind Informationen zur Verfügbarkeit, wobei vor allem der Volltext und die Onlinebestellbarkeit klar ersichtlich sein und bevorzugt werden müssen¹⁰. Ebenso müssen neben der Sortierung nach Standardrelevanz und den verbreiteten Sortiermöglichkeiten auch Kriterien wie die Verfügbarkeit von Materialien sowie die favorisierte Bibliothek sortierbar sein. Grosses Gewicht wird dem Filtern über Facetten beigemessen, da so grössere Treffermengen effizient eingeschränkt und Nulltreffermengen durch das Setzen von diesen Filtern verunmöglicht werden. Über die Oberfläche soll der Benutzer ebenfalls in der Lage sein, FRBR-Cluster nach Auflagen, Sprachen und Materialart abzurufen. Wenn immer möglich, werden online verfügbare Ressourcen eingebunden. Ob- schon vom Sinn her nicht unumstritten¹¹, ist es in SwissBib möglich, eigene Tags zu erfassen und Empfehlungen zu Treffern abzugeben. Der Export von Daten beispielsweise aus Literaturlisten in externe Dienste wie Zitierverwaltungs- und E-Learning-Systeme soll so einfach wie möglich gestaltet werden.

Aufbau und Integration

Strukturell und funktional ist SwissBib ein dreiteiliges System, bestehend aus einer Komponente zur Datenaufbereitung, dem Central Bibliographic System (CBS), einer Komponente zur Indexierung und Bearbeitung von Suchanfragen namens FAST, einer Komponente für die grafische Benutzerschnittstelle sowie einer Schnittstelle zum Zugriff auf die SwissBib-Dienste von ausserhalb namens TouchPoint. Beim Aufbau des Systems wurde ein besonderes Augenmerk auf dessen Flexibilität gelegt. So sind die beschriebenen funktionalen Komponenten auch auf der Ebene des Systems voneinander unabhängig und kommunizieren über offene Schnittstellen unter dem Einsatz offener Standards. Auf diese Weise ist es möglich, das System an zukünftige Erfordernisse anpassen zu können,

¹ Vgl. OCLC, Online Catalogs, S. 11.



Tobias Viegener

Projektkoordinator swissbib, UB Basel.

Résumé

Français

SwissBib est le futur métacatalogue réunissant les bibliothèques universitaires suisses et la Bibliothèque nationale suisse. Le but de ce projet mené par la Bibliothèque universitaire de Bâle est de réaliser un OPAC combinant les technologies du Web 2.0. La solution adoptée est un système combinant la recherche traditionnelle fédérée (*federated search*) comme celle de RERO Recherche fédérée ou le CHVK, et un système qui évite la redondance des données comme Worldcat. Les données bibliographiques des différents catalogues sont regroupées dans un «Datenpool» et régulièrement actualisées. On gagne ainsi du temps lors des recherches et une grande flexibilité dans le maniement des données. Ce système se divise en trois parties: un composant de conditionnement des données (CBS), un composant pour l'indexation et le traitement des recherches (FAST), un composant pour le graphisme (TouchPoint). Grâce à un système technologiquement à la pointe, basé sur un stockage et un traitement des données centralisées, sur l'utilisation d'une technologie de moteurs de recherche moderne et d'une interface flexible claire et structurée, ainsi que sur son intégration dans le paysage des bibliothèques suisses, le métacatalogue possède ainsi de bonnes bases pour être développé.